CRYPTANALYSIS OF THE
OIL AND VINEGAR
SIGNATURE SCHEME

AVJAD KIMIS ADI SHAMIR
WDS TECHNOLOGIES THE WEIZMANN INST
ISRAEL ISRAEL

UNIVARIATE CRYPTOGRAPHY

EXAMPLE: RSA/RABIN

\[ x^2 = v \pmod{n} \]

- Evaluation is easy
- Solution is difficult
- Can be used for encryption
- Or signature schemes
- Resisted cryptanalysis
  for 20 years

DIVARIATE CRYPTOGRAPHY

EXAMPLE: ONG-SCHNORR-SHAMIR (OSS)

\[ R = s^2 \pmod{n} \]

PUBLIC KEY  SECRET KEY  SECRET FACTORIZATION

MESSAGE: \( v \) SIGNATURE: \( x, y \)

VERIFICATION CONDITION:

\[ x^2 - k y^2 = v \pmod{n} \]

SIGNATURE GENERATION:

\[ x^2 - s^2 y^2 = v \pmod{n} \]

\[ (x+y)(x-y) = v \pmod{n} \]

\[ x+y=n \quad y=\frac{1}{2}(n+\sqrt{n}) \]

MULTIVARIATE CRYPTOGRAPHY

- Based on systems of several equations in several variables
- Most common case: Quadratic eq's:

\[ \sum_{i,j} c_{ij}^4 x_i x_j = v_i \pmod{n} \]
\[ \sum_{i,j} c_{ij}^2 x_i x_j = v_6 \pmod{n} \]

Two basic approaches:

- Few variables, large domain
- Many variables, small domain

Broken by Pollard (two different attacks on quadratic and cubic variants).
THE FIRST CASE: (FEW VARIABLES)
- SUSCEPTIBLE TO POLLARD-LIKE ATTACKS WHEN TOO MANY SOLUTIONS EXIST
- SUSCEPTIBLE TO GROEBNER-BASE ATTACKS WHEN TOO MANY EQUATIONS EXIST
- CAN BE PROVEN EQUIVALENT TO FACTORING WHEN ALGEBRAIC SYMMETRIES EXIST AND #SOLUTION IS BOUNDED
- IS NOT COMPUTATIONALLY ADVANTAGEOUS IN THIS CASE.

THE COMPLEXITY OF SOLVING SYSTEMS OF MANY ALGEBRAIC EQUATIONS IN MANY VARIABLES: FRAENKEL AND YESHA:
IT IS NP-COMPLETE EVEN WHEN:
- ALL THE EQUATIONS ARE QUADRATIC
- THE DOMAIN IS GF(2).

THE SECOND CASE: (MANY VARIABLES)
- MANY SCHEMES PROPOSED IN LAST 10 YEARS
- ALMOST ALL OF THEM WERE BROKEN
- PARTICULARLY SUSCEPTIBLE TO LOW RANK ATTACKS:

GIVEN THE QUADRATIC FORM:
\[ \sum_{i,j} c_{ij} x_i x_j \]

WE CAN (USUALLY) CHANGE IT VIA A LINEAR CHANGE OF VARIABLES TO:
\[ \sum_i d_i x_i^2 \]

AND THE RANK OF THE FORM IS THE NUMBER OF NON-ZERO \( d_i \)'S IN THIS REPRESENTATION.

PROPERTIES OF THE RANK:
- RANDOM QUADRATIC FORMS USUALLY HAVE HIGH RANK.
- IF \[ \sum_{i,j} c_{ij} x_i x_j \] HAS LOW RANK, THEN IT IS IDENTICALLY ZERO ON A LARGE LINEAR SUBSPACE.

MATRIX REPRESENTATION OF QUADRATIC FORMS: \( x^T A x \):

- \( c_{ij} \) IS TYPICALLY SYMMETRIC: \( c_{ij} = c_{ji} \)
- IN NORMAL FORM, \( A \) IS DIAGONAL.
EXAMPLE: AN UNPUBLISHED QUADRATIC SCHEME:

[CRUCIALLY DEPENDS ON PROPERTIES OF $\mathbb{GF}(2)$]

**CONSIDER THE EQUATIONS:**

\[
\begin{align*}
\gamma_1 \gamma_2 &= \psi_1 \\ 
\gamma_3 \gamma_4 &= \psi_2 \\ 
\vdots \\
\gamma_{m-1} \gamma_m &= \psi_{2m} \\
\end{align*}
\]

UNDER A LINEAR TRANSFORMATION

\[
\begin{bmatrix}
\gamma_1 \\
\gamma_2 \\
\vdots \\
\gamma_m \\
\end{bmatrix} =
\begin{bmatrix}
A \\
B \\
\vdots \\
L_{m-1,m} \\
\end{bmatrix}
\begin{bmatrix}
\psi_1 \\
\psi_2 \\
\vdots \\
\psi_{2m} \\
\end{bmatrix} \pmod{2}
\]

AND A LINEAR MIXING OF OUTPUTS

\[
\begin{bmatrix}
\psi_1 \\
\psi_2 \\
\vdots \\
\psi_{2m} \\
\end{bmatrix} =
\begin{bmatrix}
A' \\
B_1 \\
B_2 \\
\vdots \\
L_{m-1,m} \\
\end{bmatrix}
\begin{bmatrix}
\gamma_1 \\
\gamma_2 \\
\vdots \\
\gamma_m \\
\end{bmatrix} \pmod{2}
\]

**THE PUBLIC KEY:**

2m QUADRATIC FORMS $Q_i(x_1 \ldots x_n)$ in m VARIABLES $x_j$ OVER $\mathbb{GF}(2)$.

**ENCRYPTION OF BINARY CLEARTEXT:**

Evaluate the 2m $Q_1 \cdots Q_{2m}$ under $x_1 = m \cdots x_n = m$, giving binary results $\psi_1, \ldots, \psi_{2m}$.

**DECRYPTION OF BINARY CIPHERTEXT:**

$V = \psi_1 \psi_2 \cdots \psi_{2m}$

- SEEMS TO BE DIFFICULT, SINCE THE $Q_i$'S LOOK RANDOM, BUT:

- APPLY $B^{-1}$ TO CHANGE EQUATIONS TO:

\[
\begin{align*}
L_1(x_1) \cdot L_2(x_2) &= \psi_1 \\
L_3(x_3) \cdot L_4(x_4) &= \psi_2 \\
\vdots \\
L_{m-1}(x_{m-1}) \cdot L_{m}(x_m) &= \psi_{2m}
\end{align*}
\]

- SOLVE THE RESULTANT SYSTEM OF m LINEAR EQUATIONS IN m VARIABLES

---

(U NPUBLISHED) LOW-RANK ATTACK DUE TO COPPERSMITH, STERN:

\[Q = 2_1 Q_1(x) + 2_2 Q_2(x) + \cdots + 2_m Q_m(x) + L(x) \mathbb{Z}^n\]

- IN $Q$, THE COEFFICIENT OF EACH TERM $x_i x_j$ IS A LINEAR FORM IN THE NEW $z_1 \cdots z_{2m}$.

- THE LINEAR SUBSPACE WHICH MAKES $L(x) = 0$ CONTAINS HALF THE $x$ SPACE, AND MAKES $Q$ IDENTICALLY ZERO.

- LET $k = \log_2 2m$. CHOOSE A RANDOM VECTORS IN $x$ SPACE. ALL OF THEM ARE ON THIS LINEAR SUBSPACE WITH PROB

\[
\left(\frac{1}{2}\right)^k \geq \frac{1}{2m}.
\]

- FOR EACH LINEAR COMBINATION OF THESE GUESS $x$ VECTORS, THE CONDITION $Q = 0$ GIVES LINEAR EQUATION IN THE UNKNOWN $z_1 \cdots z_{2m}$.

- THE GUESSING YIELDS $2m$ LINEAR EQUATIONS IN $2m$ $x$ VARIABLES, WHICH WE CAN SOLVE.

---

THE OIL AND VINEGAR SCHEME (PATARIN, 1996)

A SIMPLE WAY TO CONSTRUCT SOLVABLE, MEDIUM RANK SYSTEM OF QUADRATIC FORMS.

\[
\begin{bmatrix}
\gamma_1 \\
\gamma_2 \\
\vdots \\
\gamma_{m} \\
\end{bmatrix} =
\begin{bmatrix}
A \\
B_1 \\
B_2 \\
\vdots \\
L_{m-1,m} \\
\end{bmatrix}
\begin{bmatrix}
\psi_1 \\
\psi_2 \\
\vdots \\
\psi_{2m} \\
\end{bmatrix} \pmod{2}
\]

\[V(x) = F(x) = \gamma_1 F_1(x) + \gamma_2 F_2(x) + \gamma_3 F_3(x) \pmod{2}
\]

PUBLISH THE COEFFICIENTS OF ALL THE $F_i(x)$ [THERE IS NO NEED TO MIX THEM LINEARLY] AS THE PUBLISHED SIGNATURE VERIFICATION KEY.
THE MESSAGE: \[ m = (m_1, \ldots, m_k) \]

THE SIGNATURE: \[ x = (x_1, \ldots, x_{2k}) \]

VERIFICATION: \[ \forall i = 1, \ldots, k: g_i(x) = m_i \]

GENERATION: CAN BE DONE EFFICIENTLY IF THE SECRET KEY A IS KNOWN:

\[ g_c(x) = m_i \Leftrightarrow f_i(y) = m_i \]

\[ f_c \circ \begin{bmatrix} 0; n \mid n; 4 \\ \hline \end{bmatrix} \Rightarrow f_c = \sum_{i,j} c_{ij} y_i y_j \text{ does not contain } y_i y_j \text{ both from first half } i, j = 1, \ldots, k. \]

DEFINITION: \( y_1, \ldots, y_k \) ARE OIL VARIABLES \( y_k+1, \ldots, y_{2k} \) ARE VINEGAR VARS.

\( f_c \) HAS ONLY OIL-VINEGAR, VINEGAR-OIL, VINEGAR-VINEGAR OCCURRENCES, SO CHOOSE ARBITRARY VINEGAR VALUES, AND SOLVE THE LINEAR SYSTEM IN OIL VARS.

MAIN PROBLEM: CAN YOU SEPARATE THE OIL AND VINEGAR VARIABLES IN THE QUADRATIC FORMS \( g_1(x) - g_2(x) \)?

---

CLAIM: ALL THE PUBLISHED QUADRATIC FORMS \( g_1(x) - g_2(x) \) ARE IDENTICALLY ZERO ON THE OIL SUBSPACE OF \( x \).

PROOF: IN \[ f_c(y) = \sum_{i,j} c_{ij} y_i y_j \]

EACH TERM HAS AT LEAST ONE INDEX IN SECOND HALF, WHICH IS 0 IN THE OIL SUBSPACE.

OBSERVATION: THE ZEROS OF EACH \( g_c(x) \) ARE USUALLY A SUPERSET OF THE OIL SUBSPACE, THEIR INTERSECTION IS USUALLY EXACTLY THE OIL SUBSPACE.

---

THE OIL SUBSPACE:

In \( y \) space: all vectors of form \( (0, \ldots, 0, 0, \ldots, 0) \)

In \( x \) space: the preimage by \( f \) of the oil \( y \) space.

THE VINEGAR SUBSPACE:

In \( y \) space: all vectors of form \( (0, \ldots, 0, x, x, \ldots, x) \)

In \( x \) space: the preimage by \( f \) of the vinegar \( y \) space.

EACH SPACE IS THE DIRECT SUM OF ITS OIL AND VINEGAR SUBSPACES.

---

BASIC IDEA: CONSIDER THE MATRIX OF COEFFICIENTS \( f_i \) OR \( g_i \) BOTH AS A QUADRATIC FORM AND AS A LINEAR MAPPING

\[ \begin{bmatrix} x_1 \cdots x_{2k} \end{bmatrix} \]

\[ g_i \begin{bmatrix} x_1 \\ \vdots \\ x_{2k} \end{bmatrix} = \text{value} \]

OR

\[ \begin{bmatrix} x_1 \\ \vdots \\ x_{2k} \end{bmatrix} \]

\[ g_i \begin{bmatrix} x_1 \\ \vdots \\ x_{2k} \end{bmatrix} = \begin{bmatrix} x_1' \\ \vdots \\ x_{2k}' \end{bmatrix} \]

CLAIM: IF \( f_i \) IS REGULAR, THEN THE LINEAR MAPPING IT REPRESENTS MAPS THE OIL SPACE ONTO THE VINEGAR SPACE, AND \( f_i' \) MAPS THE VINEGAR SPACE ONTO THE OIL SPACE.
PROBLEM: A linear change of variables \( \mathbf{y} = \mathbf{A} \mathbf{x} \) changes the quadratic form via \( \mathbf{F} \to \mathbf{A}^T \mathbf{F} \mathbf{A} \):

\[
\mathbf{F} \mathbf{x} \mathbf{x}^T = (\mathbf{A} \mathbf{x})^T \mathbf{F} (\mathbf{A} \mathbf{x}) = \mathbf{x}^T (\mathbf{A}^T \mathbf{F} \mathbf{A}) \mathbf{x}
\]

But changes the linear mapping via \( \mathbf{F} \to \mathbf{A}^T \mathbf{F} \mathbf{A} \):

\[
\mathbf{y} = \mathbf{F} \mathbf{x} \to (\mathbf{A} \mathbf{y}) = \mathbf{F} (\mathbf{A} \mathbf{x}) \to \mathbf{w} = (\mathbf{A}^T \mathbf{F} \mathbf{A}) \mathbf{x}
\]

**Corollary:** \( \mathbf{G} \) does not map the oil \( \mathbf{x} \) space to the vinegar \( \mathbf{y} \) space (as a linear mapping).

**Definition:** A linear subspace \( \mathbf{U} \) is an eigenspace of a matrix \( \mathbf{M} \) if \( \mathbf{M} \mathbf{u} = \mathbf{u} \). It is a common eigenspace of \( \mathbf{M}_1, \ldots, \mathbf{M}_t \) if \( \forall \mathbf{M} \in \mathbf{M}_1 \cup \cdots \cup \mathbf{M}_t \).

**Corollary:** Let \( T \) be the closure of all the matrices of the form \( \mathbf{G}_i^T \mathbf{G}_j \) under addition, multiplication, and multi' by a constant. Then the oil subspace is a common eigenspace of all the matrices in \( T \).

**Efficient Algorithms for Finding Common Eigenspaces**

**Definition:** Let \( p(x) \) be the characteristic polynomial of a matrix \( \mathbf{B} \).

**By Cauchy-Hamilton Theorem:** \( p(\mathbf{B}) = \mathbf{0} \).

**Lemma:** For any polynomial \( p(x) \), \( \text{kernel}(p(\mathbf{B})) \) is an eigenspace of \( \mathbf{B} \).

**Proof:** \( \mathbf{z} \in \text{kernel}(p(\mathbf{B})) \implies p(\mathbf{B}) \mathbf{z} = \mathbf{0} \).

\( \mathbf{B} \) commutes with its powers, and thus with any polynomial in \( \mathbf{B} \). So:

\[
p'(\mathbf{B}) \cdot \mathbf{B} \mathbf{z} = \mathbf{B} \cdot p'(\mathbf{B}) \mathbf{z} = \mathbf{0}
\]

So \( \mathbf{B} \mathbf{z} \in \text{kernel}(p'(\mathbf{B})) \)

The converse is not true: \( \mathbf{B} = \mathbf{I} \) the only singular polynomial in \( \mathbf{B} \) is the \( 0 \) matrix with full space as kernel.
THEOREM: If the characteristic polynomial of $B$ is irreducible, then the only eigenspaces of $B$ are $\{0\}$ and the whole space.

Proof: For each vector and linear subspace, there is a minimal polynomial in $B$ mapping it to $0$. This polynomial is a divisor of the minimal characteristic polynomial of $B$.

Let $0 \neq \mathbf{z} \in$ eigenspace $V$. $p(\mathbf{z})$ is irreducible $\Rightarrow$ monic of $\mathbf{z}$ is $p(\mathbf{z})$ itself. Thus $3, B^2, B^3, B^4, B^5, \ldots$ are linearly independent vectors, which are all in the eigenspace $V$. So $V$ is full dimensional, i.e., the whole space.

Observation: $F_i \cdot F_j = \begin{bmatrix} 0 & \ast \\ \ast & 0 \end{bmatrix}$

So its characteristic polynomials are always the products of two half degree polynomials. (This remains unchanged by similarity transformation, when $F_i \cdot F_j$ is mapped to $G_i \cdot G_j$).

Next simplest case:
The characteristic polynomial $p(\mathbf{z})$ of $B$ factors into two irreducible polynomials: $p(\mathbf{z}) = p_1(\mathbf{z}) \cdot p_2(\mathbf{z})$

Let $B_1 = p_1(B), B_2 = p_2(B), k_1 = \text{dim}(B_1), k_2 = \text{dim}(B_2)$. Then:

$K_1 \cap K_2 = \emptyset, \dim(k_1) + \dim(k_2) = n$

The only eigenspaces of $B$ are $\{0\}, k_1, k_2$, whole space.